ColonSAM: Semantic Video Segmentation

Syn Nguyen[1], Dr. Yu Tian[2]

[1]California State University - Long Beach, Syn.Nguyen01@student.csulb.edu

[2]University of Southern Florida, yu.tian2@ucf.edu

Abstract

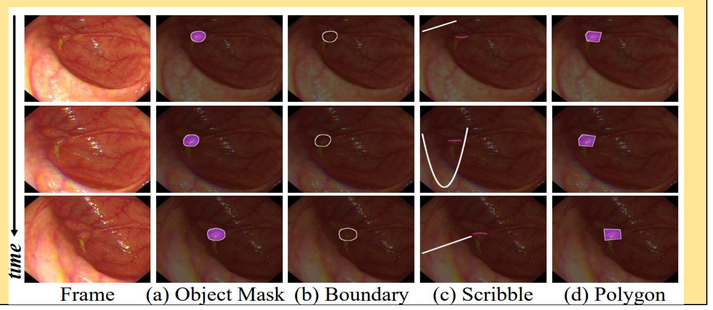
Goal:

- Develop semantic awareness segment anything model (SAM) for polyp segmentation
- Employ a zero-shot approach using point prompt optimizer agent to refine segmentation zone
- Adapt on 3D segmentation for colonoscopy dataset

Motivation:

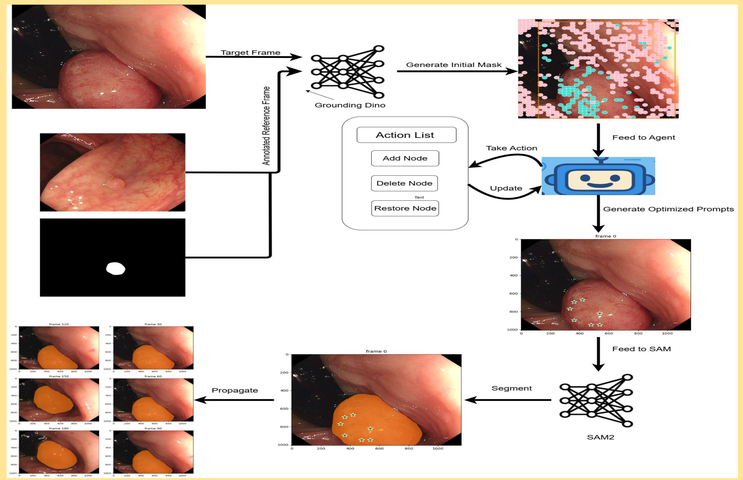
- Colorectal cancer survival rate is 91% at earlier stage (I and II). The rate is reduced to only 3% in later stage (III and IV)
- SAM2 require at least one prompt for segmentation; not automatic => not feasible to detect cancer.

Dataset:

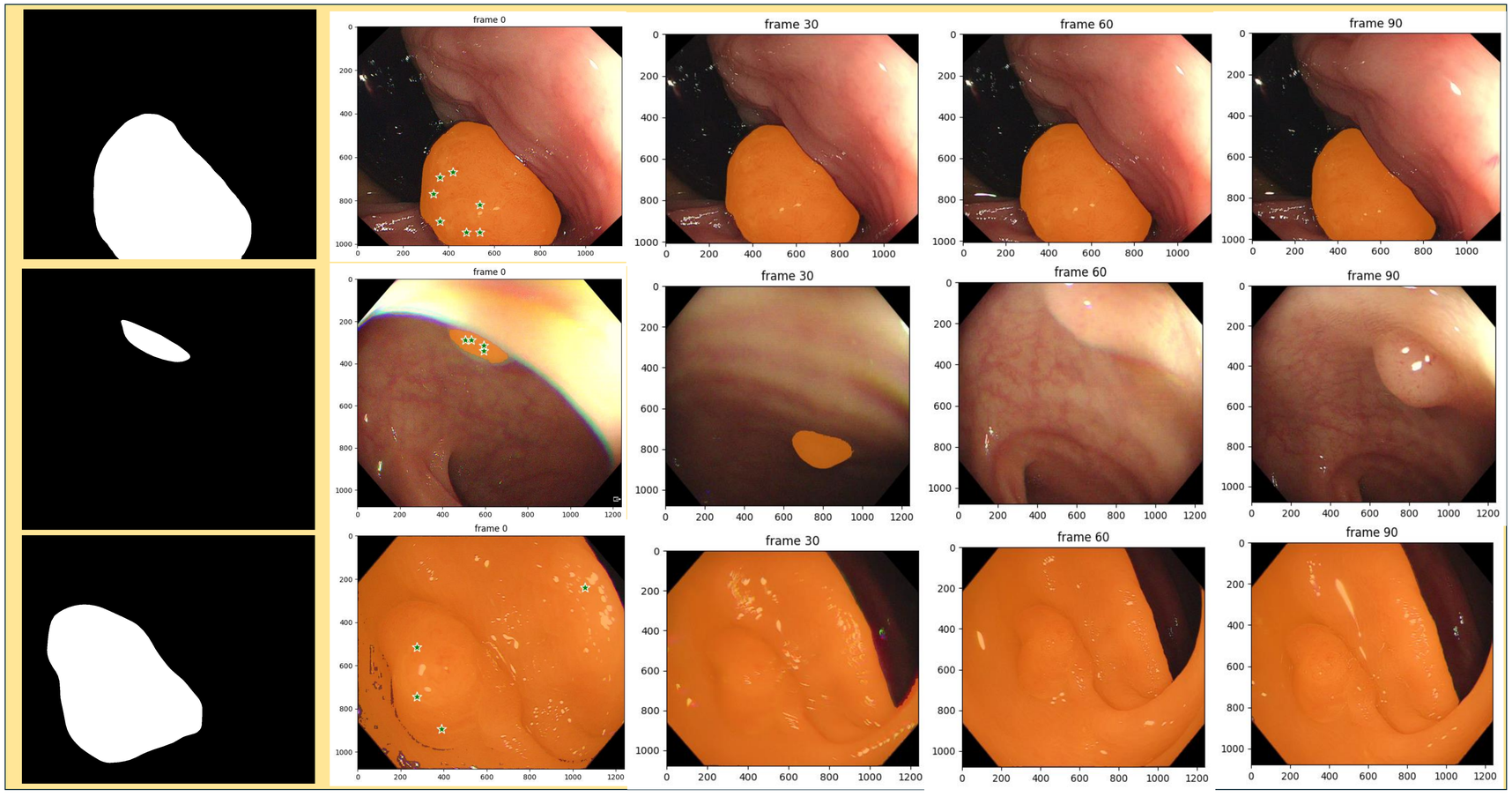


Method

- Grounding Dino for feature matching and generate initial prompt
- Train PPO agent on SUNSEG
- Utilize agent to refine initial prompts
- Generate mask with refined prompts using Segment Anything Model 2



Results



GT Frame 0 Frame30 Frame60 Frame90

Future Work

Future work:

- Fine-tune on more dataset for better generalization
- Improve model to by pass the “first-frame-problem”
- Train on more type of cancer

Challenge:

First frame problem: SAM2 require mask from the first frame to propagate to later frames. However, colonoscopy often does not have polyp in the first frame.

Reference

- **Ravi, Nikhila, et al.** "SAM 2: Segment Anything in Images and Videos." *arXiv preprint* arXiv:2408.00714, 2024, https://arxiv.org/abs/2408.00714.
- **Liu, Xueyu, et al.** "Plug-and-Play PPO: An Adaptive Point Prompt Optimizer Making SAM Greater." *2025*.
- **Ji, Ge-Peng, et al.** "Video Polyp Segmentation: A Deep Learning Perspective." *Machine Intelligence Research*, 2022.
- **Ji, Ge-Peng, et al.** "Progressively Normalized Self-Attention Network for Video Polyp Segmentation." *Proceedings of MICCAI*, 2021, pp. 142–152.
- **Misawa, Masashi, et al.** "Development of a Computer-Aided Detection System for Colonoscopy and a Publicly Accessible Large Colonoscopy Video Database (with Video)." *Gastrointestinal Endoscopy*, vol. 93, no. 4, 2021, pp. 960–967. Elsevier.
- **Biffi, Carlo, et al.** *REAL-colon Dataset*. Figshare+, 2023, https://doi.org/10.25452/figshare.plus.22202866.v2. Media.